



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

REGION IX
75 Hawthorne Street
San Francisco, CA 94105

June 22, 2000

Mr. Richard Mach
Southwest Division Naval Facilities
Engineering Command
1220 Pacific Highway
San Diego, CA 92132-5180

**SUBJECT: DRAFT VALIDATION STUDY WORK PLAN, PARCEL F
HUNTERS POINT NAVAL SHIPYARD**

Dear Mr. Mach:

The Environmental Protection Agency (EPA) has completed review of the subject document dated May 15, 2000. Our comments are included in the attachment. If you have any questions regarding these comments, please call me at (415) 744-2387.

Sincerely,

A handwritten signature in cursive script, reading "Sheryl Lauth", is written over the typed name.

Sheryl Lauth
Remedial Project Manager

cc: Mr. Chein Kao, DTSC
Mr. Brad Job, RWQCB
Mr. Jason Broederson, TTEMI
Ms. Karla Braesemle, Weston
Ms. Amy Brownell, City of SF

JUN 27 1999 15 15

OFFICE OF THE DIRECTOR

DRAFT VALIDATION STUDY WORK PLAN, PARCEL F HUNTERS POINT NAVAL SHIPYARD

General Comments:

The project team has put a lot of time and effort into this project with the resulting document as the payoff. Overall the document is acceptable with changes and should provide the data that are expected to make decisions about the low volume scenario and the Feasibility Study (FS) effort.

1. The low volume scenario, like the high volume scenario includes the areas that based on our previous data analysis are those areas that we feel should go to a FS. This document, however, is now being presented as one where through further data collections and analyses, will be suggesting that some of the area in the low volume scenario will not have to proceed to a FS. During the scoping meeting for this effort, EPA was clear that re-evaluating the low volume scenario (i.e. eliminating areas) would likely not be acceptable. We agreed that better lateral and vertical extent definition of the areas of concern may be done, however, we do not want to eliminate any area that was previously identified for evaluation as part of the FS. For example, the hope is, that we will be "...observing relationships between sediment chemistry and toxicity or bioaccumulation ..." (pC-14). The Navy is relying heavily on the ERM-Q which is a chemical distribution. The ER-M is a good predictor of an effect i.e., if a contaminant level is at or above the ER-M, there is a high likelihood of a biological effect being observed (Long et al 1995). However, the ER-M is not a predictor of effects when the chemistry data are between the ER-L and the ER-M, nor is the ER-M a good predictor of the lowest level of significant biological effect e.g., LOAEL.

One of EPA's concerns is that of using the ER-M as a metric and then assuming that the levels of concentrations at the ER-M somehow translate to "safe" levels or even "acceptable" levels when making a decision about the contaminants at Hunters Point. For instance, the ER-M for lead is 218ppm, the incidence of biological effects for concentrations above the ER-M is 90.2% (Long et al, 1995) and the calculated LC50 (survival) for *Rhepoxynius abronbius* and *Ampelisca abdita* is 170ppm (Field et al, 1996). The ER-M for fluoranthene is 5100ppb, the incidence of biological effects for concentrations above the ER-M is 92.3% (Long et al, 1995) and the calculated LC50 (survival) for *Rhepoxynius abronbius* and *Ampelisca abdita* is 3200ppb (Field et al, 1996). The ER-M for phenanthrene is 1500ppb (Long et al, 1995), the incidence of biological effects for concentrations above the ER-M is 90.3% (Long et al, 1995) and the calculated LC50 (survival) for *Rhepoxynius abronbius* and *Ampelisca abdita* is 1300ppb (Field et al, 1996).

2. The supporting material in this work plan appears to have a lot of boilerplate that is drawn from other sources. With this approach, some carry over of older documents are referenced and sometimes cited e.g., ASTM. This results in differences in what is written in these documents and what was agreed to by the agencies. This is true for counting procedures for the sediment water interface procedure and for the number of replicates for the bioaccumulation procedures. These must be updated.

Specific Comments:

1. p1, last par. Although the debris-lined shoreline areas are not of direct primary ecological concern to the intertidal areas being addressed in this effort, these areas must be addressed as a nuisance and most likely source of contaminants to Parcel F. EPA would suggest that these areas should be cleaned up by removing the debris. The evaluation of these areas is complicated by the fact that no samples can be taken from them during the confirmation effort. How will these areas be evaluated such that they may be "...found to be acting as an ongoing source of contamination to intertidal or subtidal sediments...?"
2. p2, Reference sites. EPA believes that an area like Alcatraz environs may not be appropriate because of the previous dumping activities that have taken place. Perhaps, if all of the reference sites could meet a comparison with the Water Board ambient data such that no chemical constituents was greater than the ambient number, the site could be used. Please provide a list of contaminants and concentrations for each of the reference sites.
3. p17, TIE effort. Why not use the TIE data to validate the low-volume (LV) footprint if needed? Can we assume that the Navy is taking the position that the TIE effort may provide useful information, but the overall design of the sampling for TIEs is not intended to provide information for the confirmation effort?
4. p19, Feasibility data. Please make it clear that this effort is not intended to be the final effort for a feasibility study and that when necessary, further data will be collected to complete the needs for the feasibility of options for remediation.
5. p20, Table 3-7. Data Quality Objectives for Evaluation of Sediment Dynamics. Was this table and contents discussed at a meeting with all of the Regulatory group?
6. p21, Time series measurements. This paragraph differs from the material provided on page B-18.
7. The second to last paragraph is not clear. It states that site data will be collected to be fed into the one-dimensional sediment transport model (ODSTM) and model outputs will be compared to the data collected. The same data will be used to obtain estimates from the model which will then be compared to themselves? This should be clarified.
8. p22, Table 3-8. Data Quality Objectives for Feasibility Study-Related Sediment Characterization. Were these discussed at a meeting with all of the Regulatory group?
9. p27, Last par before section 3.3. There are several locations where the hydrodynamics effort is described, please ascertain that all of these statements are consistent as some differ with others.

Appendix B.

1. pB-2, Please explain on paper the reason that the Navy believes that the five laboratory splits from a single composite sample are replicates and not sub samples. Please explain by identifying the experimental unit and what the Navy is actually testing, e.g., chemical effects, biological variance, differences in locations, etc.
2. pB-4, Table 1. For the statistical analysis, why is Dunnett's test selected rather than some

other mean separation test?

3. pB-5, Table 1. (two tables with the same designation?) Was Cd not a COPEC?
4. pB-12, Although statistical significance may be associated with ratios of sediment to tissue concentration for contaminants, this is not biological significance. For instance, the ER-M for lead is 218ppm, the incidence of biological effects for concentrations above the ER-M is 90.2% (Long et al, 1995) and the calculated LC50 (survival) for *Rhepoxynius abronbius* and *Ampelisca abdita* is 170ppm (Field et al, 1996). The ER-M for fluoranthene is 5100ppb, the incidence of biological effects for concentrations above the ER-M is 92.3% (Long et al, 1995) and the calculated LC50 (survival) for *Rhepoxynius abronbius* and *Ampelisca abdita* is 3200ppb (Field et al, 1996). The ER-M for phenanthrene is 1500ppb (Long et al, 1995), the incidence of biological effects for concentrations above the ER-M is 90.3% (Long et al, 1995) and the calculated LC50 (survival) for *Rhepoxynius abronbius* and *Ampelisca abdita* is 1300ppb (Field et al, 1996).
5. pB-20, Weight of evidence. This material has to be reworked and/or replaced.

Appendix C

1. pC-1, Last par of introduction. The second sentence is not clear as to its meaning. Why are relevant estimates of variability from both historical studies and the statistical test that will be performed required to support the determination of required sample sizes. What is the basis for stating that only chemistry data are available to accomplish the required sample sizes? Other data (Battelle), were used to evaluate bioaccumulation, are other data available for amphipods as well?
2. pC-14, Defined Strata and variation of ERM-Qs within. The strata as defined indicate a variability in ERM-Q categories which may require a point by point evaluation rather than only by these strata. Otherwise, the number of samples should be maximized to provide the greatest assurances that the strata will be adequately defined.
3. pC-15, ERM-Q threshold value (TV). How are these derived?
4. What is "w" in the formula (1)?
5. The population standard deviation (σ) is estimated from what data source? Aren't two of the three previous sampling efforts PRC, 1996 and TtEMI, 1997 the same data sets that could not be used because they were of inadequate quality for reliable estimates of variability? See ppC-11 and 14. What is the justification for using this in these exercises and not in the former?
6. pC-16, Table C-3, Sample Sizes. What is the basis of "professional judgment" for Area D? Please explain the logic for selecting these sample sizes.
7. It appears that there are a couple of arithmetic errors in the table. The 31 should be a 32 in the first column and the overall total should be 61 instead of 60.
8. pC-17, Table C-4. The hope is, "...observing relationships between sediment chemistry and toxicity or bioaccumulation ..." (pC-14). The Navy is relying heavily on the ERM-Q which is a chemical distribution. Given that the ER-M is a good predictor of an effect i.e., if a contaminant level is at or above the ER-M, there is a high likelihood of an effect

being observed (Long et al 1995). However, the ER-M is not a predictor of effects for chemistry between the ER-L and the ER-M, nor is the ER-M a good predictor of the lowest level of significant biological effect e.g., LOAEL. The information provided in the table below shows the percent biological effects for contaminant concentrations above the ER-M. For those contaminants where ER-Ms are available, the percent effect is significant. The lowest percentages are for those contaminants that have little concern for acute effects, but are very much a concern for bioaccumulative effects e.g., mercury, organochlorine and PCBs. The bottom line is that ER-Ms cannot be considered as protective when viewed from a toxicity angle and the data from Long et al (1995) shows this to be true.

Contaminant and percent incidence of biological effects for concentrations above ER-M¹

	%Effect Above		%Effect Above		%Effect Above
<u>Metals</u>	<u>ER-M</u>	<u>PAHs</u>	<u>ER-M</u>	<u>Pesticides , PCBs, and Organometals</u>	<u>ER-M</u>
Arsenic	63%	Anthracene	85.2%	alpha-Chlordane	(NA)
Chromium	95%	Benzo(a)anthracene	92.6%	4,4'-DDD	(NA)
Copper	83.7%	Benzo(a)pyrene	80%	4,4'-DDE	(NA)
Lead	90.2%	Benzo(b)fluranthene	(NA) ²	4,4'-DDT	(NA)
Mercury	42.3%	Benzo(g,h,i)perylene	(NA)	Total PCBs	51.0%
Nickel	16.9%	Benzo(k)fluoranthene	(NA)	Tributyltin	(NA)
Silver	92.8%	Chrysene	88.5%	Total DDT	53.6%
Zinc	69.8%	Dibenz(a,h)anthracene	66.7%		
		Fluranthene	92.3%		
		Indeno(1,2,3-cd)pyrene	(NA)		
		Naphthalene	88.9%		
		Phenanthrene	90.3%		
		Pyrene	87.5		

¹Data from Long et al, 1995.

²Not Available.

EPA is requesting that the Navy provide a plot of each of the contaminants within each of the strata using the data that comprised the ERM-Qs. Perhaps the cumulative plot (Fig C-11, pC-10) could be used to show the distribution of each ERM-Q category.

Appendix D.

1. p20, Sample Processing. pH should be measured in the field.

Appendix E.

1. pE-39, first bullet on page. How many reps are planned, two or three and how many for depuration?
2. pE-59, Table E-7. Were these DQOs discussed with the regulatory group?
3. pE-61, Table E-8. Were these DQOs discussed with the regulatory group?

4. pE-62, Table E-9. First row, third column, Assessment. Isn't one of the goals, exposure response gradient? Should this be listed here?
5. pE-63, Table E-10. Should the benchmark/thresholds be ER-M instead of ER-L?

Citations

Long, E.R., D.D. Macdonald, S.L. Smith, and F.D. Calder. 1995. Incidence of Adverse Biological Effects Within Ranges of Chemical Concentrations in Marine and Estuarine Sediments. Environmental Management Vol. 19, No. 1, pp81-97.

Field, L.J., D.D. MacDonald, C.G. Severn. 1996. Use of a Sediment Toxicity Database for Evaluating Matching Sediment Chemistry and Toxicity Data. 1996. HAZMAT Report 97-1.